

Multiple Choice

1. If we observe the timing of a stochastic process to be between 30 and 36 seconds with a confidence of 90%, what is the most likely new confidence interval if you are asked for 95% confidence?
 - a. Between 25 and 41 seconds
 - b. Between 32 and 34 seconds
 - c. Between 30 and 36 seconds
2. Suppose there are two fast-food coffee chains. The average end-to-end coffee preparation time at Java James is 4 minutes and 52 seconds. Koehler's Kafe, on the other hand, averages 5 minutes and 5 seconds. Customers expect to wait no more than 5 minutes for coffee. Which statement can we **conclusively** make? (Select all that apply)
 - a. James has more customers served within 5 minutes
 - b. James has less variability in his coffee-making process
 - c. There is more value-added time in Koehler's process
 - d. None of the above
3. What advantages does single-variate linear regression have over correlation? (Select all that apply)
 - a. Predicts the direction change in Y, given a directional change in X
 - b. Predicts the magnitude change in Y, given a magnitude change in X
 - c. Predicts Y, given a value of zero for X (the intercept)
 - d. Implies that the change in Y was caused by the change in X

Short Answer

4. Please describe a project where you applied your combination of analytical and business skills to a business problem, which lead to a positive and tangible impact. Your response should address the following questions:
 - i. What was the problem, how did you identify it, and who was the customer?
 - ii. How did you acquire the data, and what uniqueness or complexity made it challenging?
 - iii. What analytical or modeling steps did you subsequently take and why did you choose them?
 - iv. Describe the business benefit attained from your work. Were there any additional risks or effects as a result of the proposed change?
 - v. Reflecting on your project, what could you have done better?
5. You are asked to lead an analysis to forecast the busiest hours (e.g. 6pm) in pharmacy, first by generating the expected hourly "demand curve" for each of our stores in 2018. A snippet of the available data is below.

Submit a query to generate this desired output, feel free to use any SQL syntax/implementation (e.g. SQL Server, DB2) you are most comfortable with. What assumptions did you make, and what clarifications might you ask for?

SELECT TOP 5 * FROM prescription_create_audit;

store_number	prescription_number	date_time_prescription_created
123	265766724	3/19/2018 14:36
123	818098296	9/28/2018 10:03
456	75802089	8/23/2018 8:52
123	773736822	10/17/2019 16:45
456	794357442	11/14/2018 17:33
		(etc...)